Citation: Ortega, A., Geraldi, N. R., Díaz-Rúa, R., Bachmann Ørberg, S., Wesselmann, M., Krause-Jensen, D., & Duarte, C. M. (2020). A DNA mini-barcode for marine macrophytes. Molecular Ecology Resources. doi:10.1111/1755-0998.13164

A DNA mini-barcode for marine macrophytes

Alejandra Ortega^{1*}, Nathan R. Geraldi², Rubén Díaz-Rúa², Sarah Bachmann Ørberg³,
Marlene Wesselmann⁴, Dorte Krause-Jensen³, and Carlos M. Duarte^{2,5}

^{1.} Biological and Environmental Science & Engineering Division (BESE), King Abdullah
University of Science and Technology (KAUST), Thuwal, Saudi Arabia

^{2.} Red Sea Research Center (RSRC), King Abdullah University of Science and Technology
(KAUST), Thuwal, Saudi Arabia

^{3.} Department of Bioscience, and Arctic Research Centre, Aarhus University, Denmark

^{4.} Global Change Research Group, Institut Mediterrani d'Estudis Avançats IMEDEA (CSIC-
UIB), Esporles, Spain

^{5.} Computational Bioscience Research Center (CBRC), King Abdullah University of Science
and Technology (KAUST), Thuwal, Saudi Arabia

*Correspondence: Alejandra Ortega, email: alejandra.ortega@kaust.edu.sa

Abstract

Studies focusing on marine macrophyte metabarcoding from environmental samples are scarce, due to the lack of a universal barcode for these taxa, and to their poor representation in DNA databases. Here, we searched for a short barcode able to identify marine macrophytes from tissue samples; then, we created a DNA reference library which was used to identify macrophytes in eDNA from coastal sediments. Barcoding of seagrasses, mangroves and marine macroalgae (Chlorophyta, Rhodophyta and Phaeophyceae) was tested using 18 primer pairs from six barcoding genes: the plant barcodes *rbcL*, *matK* and *trnL*, plus the genes ITS2, COI and 18S. The 18S gene showed the highest universality among marine macrophytes, amplifying 95-100% of samples; amplification performance of the other barcodes was limited. Taxonomy was assigned using a phylogeny-based approach to create an 18S DNA reference library. Macrophyte tissue sequences were accurately identified within their phyla (88%), order (76%), genus (71%) and species (23%). Nevertheless, out of 86 macrophytes tested, only 48% and 15% had a reference sequence at genus and at species level, respectively. Identification at these levels can be improved by more inclusive reference libraries. Using the 18S mini-barcode and the reference library, we recovered eDNA from 21 marine macrophytes in sediments, demonstrating the barcode's ability to trace primary producers that contribute to blue carbon. We expect this barcode to also be useful for other ecological questions, such as tracing macro primary producers in marine food webs.

Keywords: DNA barcoding, mini-barcode, macroalgae, marine macrophytes, environmental DNA, eDNA.

Introduction

Marine macrophytes, especially macroalgae, form highly productive ecosystems, with a global net primary production of about 2 Pg C year⁻¹ spread across 4 million km² in coastal areas (Duarte, 2017). Marine macrophytes export most of their net production to the global ocean and seafloor (Duarte & Cebrián, 1996; Duarte & Krause-Jensen, 2017; Krause-Jensen & Duarte, 2016; Ortega et al., 2019). Exported macrophyte biomass play important roles in marine ecosystems, such as contributing to organic carbon sequestration in coastal sediments and in the deep-sea (Krause-Jensen & Duarte, 2016; Ortega et al., 2019), and supporting food webs beyond the macrophyte habitat (Baker et al., 2018; Queirós et al., 2019). However, tracing the export of marine macrophytes is challenging (Geraldi et al., 2019). Several techniques have been used to determine the origin of primary producers in studies of marine sediments, food webs and diet analyses based on gut content; among these, stable isotopes ¹³C and ¹⁵N in bulk biomass are commonly used (Loneragan, Bunn, & Kellaway, 1997; Raven et al., 2002; Shimada et al., 2014). Nevertheless, isotopic signals often overlap between primary producers, failing to provide conclusive fingerprinting of individual groups of primary producers (Geraldi et al., 2019; Loneragan et al., 1997; McKee, Rogers, & Saintilan, 2012; Raven et al., 2002).

Environmental DNA (eDNA) metabarcoding could enhance our ability to identify marine macrophytes. Extraction and sequencing of eDNA —or the DNA left by organisms in their surroundings, is increasingly used to assess and monitor marine biodiversity (Aylagas, Borja, Muxika, & Rodríguez-Ezpeleta, 2018; Stat et al., 2017; Thomsen et al., 2012; Thomsen & Willerslev, 2015). However, marine macrophytes are underrepresented in the eDNA assessments from marine samples with the first studies only recently appearing. A pioneering study used eDNA to resolve seagrass, mangroves, and saltmarshes in sediment samples, but did not identify

macroalgae (Reef et al., 2017). One study reported marine biodiversity from eDNA using ToL-metabarcoding, and although four photosynthetic eukaryote phyla were recovered, the results did not provide deeper differentiation among major macrophyte lineages (Stat et al., 2017). Another study recovered macroalgal eDNA from sediments, but only identified four species while the remaining were classified as OTU (Queirós et al., 2019).

In contrast to the universality of barcoding in animals (Hebert, Ratnasingham, & de Waard, 2003), there is no universal barcoding for plants nor for marine macrophytes (CBOL, 2009; Hebert, Gregory, & Savolainen, 2005; Hollingsworth, Graham, & Little, 2011; Taberlet et al., 2007). The complex origin and broad genetic divergence of macrophytes challenge the development of a single barcode for these taxa (Hollingsworth et al., 2011; Pennisi, 2007; Rubinoff, Cameron, & Will, 2006). Eukaryota has several supergroups, and marine macrophytes belong to two of them (Burki, Roger, Brown, & Simpson, 2019). Embryophyta (plants), Rhodophyta and Chlorophyta are clades in the Archaeplastida supergroup, and Phaeophyceae is a whole clade in the TSAR supergroup. On the contrary, Metazoa, the target of many eDNA studies to date, is only a single clade within the supergroup Amorphea (Burki et al., 2019). Therefore, barcode development for the widely diverse photosynthetic eukaryotes, and in particularly understudied groups such as marine macrophytes, represents a higher complexity in comparison to the development of barcodes for animals (CBOL, 2009; Chase et al., 2007; Hebert et al., 2003; Hollingsworth et al., 2011; Pennisi, 2007; Rubinoff et al., 2006).

Many of the available barcodes and primers for photosynthetic eukaryotes are designed for Embryophyta, mainly angiosperms, and are not recoverable in other groups such as marine macrophytes, especially macroalgae (CBOL, 2009; Chase et al., 2007; De Groot et al., 2011; Kuo, Li, Chiou, & Wang, 2011; Lahaye et al., 2008; Taberlet et al., 2007; Yao et al., 2010; Yu, Xue, &

Zhou, 2011). With the exception of some COI primer combinations that can identify taxa within Rhodophyta (Saunders, 2005), regions such as trnH-psbA, trnL, matK, and rbcL have been tested in selective macroalgal species with little PCR or sequencing success (Du et al., 2014; Heinecke et al., 2017; McDevit & Saunders, 2009; Saunders & Kucera, 2010; Saunders & McDevit, 2012).

The limitations for marine macrophyte differentiation are likely due to the minimum scientific interest in studying these groups, reflected in the insufficient availability of DNA resources such as barcodes and reference libraries. This lack of DNA resources restricts the application of eDNA analysis to fingerprint marine macrophytes from the environment (Geraldi et al., 2019; Reef et al., 2017). Moreover, eDNA analyses require barcodes able to recover short DNA amplicons from environmental samples.

Although quality of eDNA depends on several abiotic factors (Barnes et al., 2014), eDNA quality is generally lower than the quality of DNA isolated directly from the organisms (Hajibabaei et al., 2006; Little, 2014); therefore, metabarcoding from environmental samples relies mainly on fragmented eDNA (Epp et al., 2012). The average size of amplicons from most available plant barcodes is over 500 bp, whereas much shorter barcodes (100-300 bp) are required for amplification of fragmented eDNA (Epp et al., 2012; Hajibabaei et al., 2006; Little, 2014). Still, the level of taxonomical resolution reduces with barcode length. This reduction may be an acceptable trade-off for environmental applications requiring biological assessment of broad taxonomic categories, where classification of marine macrophytes at the level of major lineages would suffice. For our purpose of differentiating macro primary producers in coastal sediments, the identification of marine macrophytes at family, order and even phylum level is an acceptable trade-off, as these levels of taxonomical identification are not achieved with traditional methods such as the stable isotopes.

In this study, we searched for barcodes and primer pairs able to fingerprint green, brown and red macroalgae (Chlorophyta, Phaeophyceae and Rhodophyta, respectively), and seagrasses and mangroves (Embryophyta). We created a DNA reference library and tested the identification of marine macrophytes from eDNA collected in coastal sediments, using the chosen primer pair. We searched both barcode universality and barcode specificity for each macrophyte group. The tested barcodes include the 18S gene, widely known for its universality across eukaryotes (Amaral-Zettler, McCliment, Ducklow, & Huse, 2009; Guardiola et al., 2015; Stoeck et al., 2010), and the CBOL suggested regions *rbcL* (Little, 2014; Reef et al., 2017), *matK* (CBOL, 2009), *trnL* (Epp et al., 2012; Taberlet et al., 2007), ITS2 (Yao et al., 2010) and COI (Saunders, 2005).

Materials and Methods

Primer selection

We tested the amplification of barcodes within *rbcL*, *matK*, *trnL*, COI, ITS2 and 18S regions using 18 primer pairs on DNA extracted from tissue of macroalgae, mangroves, seagrasses, and terrestrial plants (including ferns, gymnosperms and angiosperms). Plastid regions *rbcL*, *matK* and *trnL* are recommended candidates by the CBOL Plant Working Group (CBOL, 2009). *rbcL* and *matK* genes present high levels of species discrimination in angiosperms (Lahaye et al., 2008; Reef et al., 2017); *rbcL* has also been used for identification of marine angiosperms in sediment samples (Reef et al., 2017). Although *trnL* intron and the nuclear internal transcribed spacer ITS2 have low phylogenetic resolution in closely related samples, these regions have identified less studied groups such as bryophytes and gymnosperms (Taberlet et al., 2007). Furthermore, these barcodes have been reported to perform better than others for degraded plant samples (Epp et al., 2012; Taberlet et al., 2007; Yao et al., 2010).

The initial aim was to find a specific barcode for each marine macrophyte group. Nevertheless, none of the primer pairs were lineage-specific. Since the proposed combinations of plastid genes were less efficient beyond angiosperms, and since our needs do not require high taxonomical identification and can be fulfilled with resolution at even phylum level, we included the universal barcodes from the mitochondrial COI and the nuclear ribosomal 18S genes. Both regions have successfully identified some macroalgal groups, although based on long amplicons (McDevit & Saunders, 2009; Queirós et al., 2019; Saunders, 2005).

Of the 18 primer pairs tested, five were designed in this study, while the rest were selected from the literature (Table 1). EcoPrimers software (Riaz et al., 2011) was used to design the primers from regions *rbcL* and *matK*, targeting the taxa Embryophyta (taxid: 3193), Chlorophyta (taxid: 3041), Rhodophyta (taxid: 2763) and Phaeophyceae (taxid: 2870) available in the NCBI database (<https://www.ncbi.nlm.nih.gov/>). Targeted amplicons were between 120 and 220 bp length, following a mini-barcode approach (Little, 2014); no primer mismatches were allowed.

Selection of the primers was based on their ability to amplify the majority of species across a wide range of macrophytes. Amplification of the 18 primers was initially tested on 10-20 species, including green, red, and brown macroalgae, and land and marine embryophytes (Table 2). Based on their performance (that is, amplification of at least one sample per lineage), we selected some primers for a second test on a broader range of samples. For some primers, the test was stopped when the amplification was inconsistent with the new set of tissue samples, while for others, the number of samples was increased when the primer was consistently amplifying most or all the species tested (Table 3, Figure 1).

Primer amplification and taxonomic resolution were tested *in silico* using EcoPCR software (Ficetola et al., 2010), focusing on sequences from the targeted macrophyte lineages

(Embryophyta, Chlorophyta, Phaeophyceae and Rhodophyta) available in the EMBL database (<https://www.embl.de/>), with eight mismatches, and amplicon length between 120 and 220 bp.

Sample collection

To create a DNA reference database, tissue of 126 macrophytes and terrestrial plants were collected along the Arabian coast of the Red Sea, in Cyprus, in Finland, and in Greenland. These samples included macrophytes from lineages of interest for this research: terrestrial and marine angiosperms (seagrass and mangrove) and macroalgae (within Chlorophyta, Phaeophyceae and Rhodophyta). Seven gymnosperms and two ferns were included to test the primer universality for macrophytes (Table 2). Seagrasses and macroalgae were harvested by tearing the whole individual from the substrate, sometimes requiring snorkeling or diving (1-5 m), while several leaves were collected from mangrove and land plants; each sample was placed into a zip-lock bag and frozen or dried as soon as possible until DNA extraction. Seagrasses, mangroves and other Embryophyta were identified at species level following common botany knowledge, while macroalgae were identified at species or genus level following guides (Lieske, Fiedler, & Myers, 2004; Pedersen, 2011; Richmond, 2002), or the AlgaeBase (Guiry, 2013) when the species were not available in the taxonomic keys.

To demonstrate the utility of the barcode, environmental DNA samples were collected from surface sediments (top 1 cm) of five seagrass meadows and five mangrove forests along the Arabian coast of the Red Sea (28°-18°N, 35°-41° E; Figure 3). Seagrass species found in the meadows during sampling included *Halodule uninervis*, *Halophila stipulacea*, *H. decipiens*, *H. ovalis*, *Thalassodendron ciliatum* and *Thalassia hemprichii*; while the mangrove *Avicennia marina* was the only species found in the mangrove forests. In each sampling site, five sediment

replicates were collected randomly along the meadow or the forest, then pooled together as a single eDNA sample. Sediments were collected using gloves and bleach-sterilized 50 ml tubes, placed into sterilized bags, kept cold and protected from the light, then frozen at -80 °C until eDNA extraction.

DNA extraction

To isolate DNA from the targeted marine macrophyte, we removed epiphytes growing on the surface. The tissue surface was cleaned with a scalpel, and sterilized by applying bleach and ethanol for a fixed time (Coombs & Franco, 2003). However, we noticed after sterilization that the procedure should be adjusted to each tissue sample, as the bleach removed pigments and it has been shown to remove chloroplast DNA (Aires, Marbà, Serrao, Duarte, & Arnaud-Haond, 2012). For instance, green algae (i.e. *Ulva lactuca*) should be bleached for a shorter time (1 min) than brown algae (i.e. *Turbinaria ornata*; 3 min), whose cell wall have greater resistance to bleach-sterilization. Homogenization was done grinding the tissue with a pestle in a mortar with liquid nitrogen. The pestle and mortar were cleaned between samples with 5% bleach and 70% ethanol. Genomic DNA was extracted from 100 mg of homogenized tissue using the NucleoSpin Plant II kit (Macherey-Nagel) following the manufacturer's protocol.

Intracellular and extracellular environmental DNA from each of the 10 sediment samples were extracted with chloroform:isoamyl alcohol following the method developed by Lever et al. (2015). Extracellular eDNA is the pool of DNA that can be recovered directly from the sediment matrix, while intracellular eDNA is the DNA within cells, that will become available in the sediment matrix after cell lysis (Lever et al., 2015). A step-by-step protocol of this method is in the Supplementary Information.

DNA amplification and sequencing

To avoid cross-contamination, DNA extractions and PCR preparations were performed in separate areas of the laboratory and with separate material. Benches and instruments were frequently cleaned either with 5% bleach and 70% ethanol or using RNaseZap (Invitrogen) during the process. The PCR station and consumables were UV-sterilized before use. Annealing temperature per primer was estimated after running a triplicated gradient PCR (40-60 °C, optimal T_m are reported in Table 1). PCR cocktails of each primer were performed in five replicates of 10 µl reactions, by adding 5 µl of QIAGEN Multiplex PCR Master Mix (which does not contain a proofreading polymerase), 1 µl of each primer (1.0 µM), 1 µl of DNA template, and 2 µl of PCR water. Thermocycler program for primer 18S-V4 was as described in Stoeck et al. (2010), while for the other primers was: 15 min at 95 °C, followed by 35 cycles of 30 s at 94 °C, then 45 s at each primer T_m (see Table 1), 90 s at 72 °C, and a final extension at 72 °C for 10 min. The same program was used for environmental DNA. To reduce PCR inhibitors and improve amplification performance, eDNA samples were diluted at either 1:1, 1:10, 1:50 or 1:100 before preparing PCR reactions. Blanks –reagents with DNA-free water instead of sample– for DNA/eDNA extractions and for PCR were prepared and run for each primer. Mangrove *Avicennia marina* was included as PCR positive control for all primers. PCR products were visualized on 2% agarose gel. Successful amplicons from the best two primer pair candidates were prepared for sequencing by excising amplicons from gel, purifying with Wizard SV Gel and PCR Clean-up purification kit (Promega). DNA concentration was measured by Qubit 2.0 Fluorometer (Invitrogen).

Since we used up most of the PCR amplicons of one of the primer pair candidates (18S2) during the initial sequencing tests that failed, we could not create a DNA reference library that included

all collected species for the two candidate primers. Thus, eDNA sequencing was performed only with one primer pair (Euka02). DNA from tissue samples and eDNA from sediments were sequenced, separately, using the MiSeq Illumina platform. Amplicons were cleaned and indexed (Nextera XT Index Kit) following Illumina Metagenomic Sequencing Library Preparation protocol (Illumina). Average sequencing depth was 92,662 (SD $\pm$ 33,771) reads per sample.

Before opting for Miseq of the macrophyte tissue, we initially Sanger sequenced DNA from non-sterilized tissue using primers Euka02, 18S1, 18S2, minirbcL, COI GF1-DR1, sgmK1 and sgmK2. Those sequencing attempts failed. For most land plants, there was no consensus on the forward and reverse sequences, and blast was biased towards well-represented species in the databases (Supplementary Table 1). For marine macrophyte, the poor-quality of the sequences was likely influenced by additional DNA sources from organisms growing on the surface of the tissue. This was confirmed from the MiSeq sequencing which included epiphytic fauna (sponges, anemones, mollusks). In addition, some Sanger sequences had multiple nucleotides per site and did not blast with any taxa. We attempted Sanger sequencing a second time with amplicons from sterilized plant tissue using primers 18S2 and minirbcL. However, sequencing attempts failed again (Supplementary Table 1).

Sequencing analyses and clustering

Illumina MiSeq sequencing outputs were demultiplexed using the Illumina protocol and primers were trimmed using Cutadapt 1.17 with default settings (Martin, 2011). FASTQ files were analyzed following default DADA2 pipeline version 1.8.0 (Callahan et al., 2016) using R Studio version 1.1.453 (Team, 2015); sequence variants were generated. DADA2 inspects the read quality, preprocesses the data by filtering and trimming the sequences (parameters as default,

TruncLen=c(110,110)), learns the error rates, eliminates redundant reads, merges paired reads, removes chimeras and then provides an amplicon sequence variants table.

DNA sequences from macrophyte tissue were compared against the SILVA database using the assignTaxonomy function in DADA2, but the taxonomical assignment was unsuccessful. Assigning species to sequences using reference databases depends on availability of reference sequences, yet most macroalgae taxonomy is undescribed and we found that there were not enough genome reference databases, thus we could not determine which species was successfully sequenced. As the traditional assignment techniques preclude identification of most species of macroalgae, we opted for a phylogeny-based approach for taxonomical assignment. The top five most abundant MiSeq sequence variants per tissue sample were aligned (MUSCLE), along reference sequences from a costume database of 102 macrophytes from similar species, genera, or orders. This database was based on the SILVA 132 SSURef_Nr99_tax reference database (<http://www.arb-silva.de>), but included only sequences that amplified *in silico* using the virtualPCR function with default parameters from the Insect package (Wilkinson, Davy, Bunce, & Stat, 2018) in R. The sequences were clustered in GENEIOUS version 11.1.4 (Biomatters Ltd.) with a PhyML tree using TN93 model and 1,000 replicates (Guindon et al., 2010). Sequence clusters were edited in FigTree (Rambaut, 2007). From the five most abundant sequence variants, those that clustered near the expected reference taxon or the closest available (genus, family, order) were chosen to be the representative sequence of the species sampled, and a DNA reference library of these macrophytes was created. In total, the created reference library included 88 sequence variants from all the 86 tissue samples. Using this DNA reference library and the SILVA database, taxonomy was assigned to the eDNA sequences extracted from sediments using the assignTaxonomy function in DADA2.

Results

Primer performance

The nuclear ribosomal 18S barcode performed the best. The four 18S primers amplified across all macrophytes (Table 3). However, two primers (18S1 and 18S-V4) amplified only 30-67% of the macrophytes, and an even lower percentage of macroalgae (Supplementary Table 2). Conversely, primers 18S2 (V9 region) and Euka02 (V7 region) amplified well in all lineages (95-100%, 88-119 species; Table 3). Both Euka02 and 18S2 amplified *in silico* the highest number of macroalgae (37 and 29 macroalgae, respectively). Hence, Euka02 and 18S2 were chosen for the sequencing tests.

Performance of the primer pairs from the nuclear internal transcribed spacer ITS2 and plastid regions trnL, matK and rbcL was limited for macroalgae (Table 3; Supplementary Table 2). All primers from trnL and three of the four matK primers amplified angiosperms but failed to amplify macroalgae; overall amplification was between 21 and 29% of the species tested. The remaining matK primer pair (sgmatK2) amplified all lineages (83%, 35 out of 41 species), although macroalgae amplification was limited to 57% of species tested (8 out of 14 species); this primer did not amplify *in silico* any of our macroalgae (Supplementary Table 4). One of the rbcL primers amplified in all lineages but did not work in all the tissue samples (60%, 9 out of 15 samples); *in silico* amplification of macroalgae was restricted to four samples (Supplementary Table 4). Although the minirbcL primer pair amplified *in silico* all Embryophyta, none of the land angiosperms were amplified experimentally; this primer amplified 47% of the other macrophytes (14 out of 30 tissue samples; Table 3, Supplementary Table 2).

Whereas *in silico* amplification of algae was higher in primers that poorly amplified our macroalgae samples, most of those *in silico* amplifications correspond to unicellular microalgae (Supplementary Table 3). For instance, primer minirbcL amplified *in silico* 5,739 algae but only three of those are macroalgae tested in this study. COI GF1-DR1 primer amplified 2,125 algal species *in silico*, 40 of which belong to our macroalgae; nevertheless, we tested amplification of 19 of those species and only three macroalgae amplified experimentally (Supplementary Table 4). This result was similar for the remaining COI primers, which only amplified a small percentage of macroalgae (2-11%; Table 3; Supplementary Table 2). Moreover, none of the COI primers amplified any of our Embryophyta neither *in silico* nor experimentally. On the contrary, 37 out of 214 algae that amplified *in silico* using primer Euka02 belong to this study (Supplementary Table 4). Furthermore, 41 and 29 macroalgae amplified experimentally but did not amplify *in silico* with the primers 18S2 and Euka02, respectively (Supplementary Table 4).

Amplification of environmental DNA from seagrass and mangrove sediments was tested with the primers sgmK2, the two rbcL primers, and with 18S2 and Euka02 from the 18S barcode; all primers amplified with dilutions of either 1:1, 1:10, 1:50 or 1:100. The best performance was achieved with the 18S barcodes.

DNA reference library

MiSeq sequencing from macrophyte tissue was conducted for amplicons from primers 18S2 (11 samples) and Euka02 (86 samples). For the DNA reference library, we focused on the sequences from the Euka02 primer for which we had the highest number of samples. Similarly, as the DNA reference library was a requirement for macrophyte fingerprinting, eDNA metabarcoding of sediments was performed only on Euka02 amplicons. Nevertheless, we have successfully used

both primers for different applications on eDNA from marine sediments (unpublished data, see amplification in Supplementary Figure 2).

Taxonomical assignment based on blasting against the SILVA database was ineffective for most marine macrophytes, due to a lack of existing DNA references for those species in the database. Thus, we followed a phylogeny-based approach and clustered the sequences from our 86 macrophytes along with 102 reference sequences; these reference sequences ranged from the same species or genus, when available, to other species within the same order of the tested macrophytes.

Since the MiSeq output for each single macrophyte sample generated several sequence variants, we chose the five most abundant sequence variants per tissue sample. These sequences represented about 88% of the total reads per sample (94% in Embryophyta, 93% in Phaeophyceae, 85% in Rhodophyta, and 74% in Chlorophyta). We assumed that the actual macrophyte is within the top sequence variants, while the others could be epiphytes growing on the surface of the tissue or endophytes living within the macrophyte cells. Taxonomy was assigned to the sequence variant (or sequence variants) that clustered near the expected reference taxon. Following this approach, we created a DNA reference library for 76 out of the 86 macrophytes sequenced with the Euka02 primer.

Discrimination success

Barcoding with the 18S gene clearly distinguished each lineage of macrophytes. Maximum-likelihood clusters from the Euka02 primer displayed high taxonomical cohesion, except for green algae (Figure 2; similar findings from the primer 18S2 are shown in the Supplementary Figure 1).

335 Sequences from embryophytes (marine angiosperms and land plants), brown and red algae
336 grouped in their own assemblages and clustered with the expected reference sequences (Figure 2).

337 Barcoding using the primer Euka02 accurately assigned 88% of macrophytes to their lineage:
338 Embryophyta 95% (19 out of 20 species), Rhodophyta 81% (17 out of 21 species), and
339 Phaeophyceae 86% (24 out of 28 species; Table 4). Although no Chlorophyta clustered with the
340 reference sequences, most green algae (94%, 16 out of 17 species) formed their own clade apart
341 from the other phyla. Order level was accurately aligned to 77% of the samples, while 52% of
342 samples were aligned to the correct family (Table 4). Genus level was accurately aligned to 79%
343 (37 out of 47) of the samples for which a reference genus was available. Four macrophytes
344 clustered with their reference at species level, although only 15 out of 86 macrophytes had a
345 reference sequence at this taxonomic rank available in the reference database.

346 All Embryophyta formed a single clade and the barcoding region could separate angiosperms
347 from gymnosperms, and monocots from eudicots. Exceptions were the seagrasses *Thalassia*
348 *hemprichii* that clustered within the red algae, and one of the two samples of *Posidonia oceanica*
349 that, although classified within the phylum, clustered with gymnosperms (Figure 2); the reference
350 sequences of these species clustered in the correct clade. *Cupressus sempervirens*, the sole
351 gymnosperm sequenced here, accurately clustered with the reference species *C. tonkinensis* and
352 with *Pinus taeda*, which belongs to the same order (Pinales). The seagrass order (Alismatales)
353 formed a subclade clustering most samples with all reference sequences for this order. However,
354 the seagrass family Hydrocharitaceae clustered in several branches while all reference sequences
355 from the family Cymodoceae clustered together. *Cyperus* sp. wrongly clustered within
356 Alismatales, although both taxa are monocotyledons. The three mangroves samples clustered at

species level, and the land plants clustered at genus level — no species reference were available for land plants.

The Rhodophyta lineage clustered together and displayed two distinctive branches. The larger branch included only species of the subphylum Eurhodophytina, while species from several subphyla were included in the smaller branch (Figure 2). Prior to sequencing, we morphologically identified each collected species. However, we were uncertain of accurate identification of a few Red Sea macroalga species due to a lack of identification keys. The main uncertainties in Rhodophyta were species of the *Laurencia* complex, a genus that was recently split in four genera (Martin-Lescanne et al., 2010). We were unsure of four samples in this complex: we identified three of them as *Laurencia* species, while one was morphologically identified as *Chondrophycus*. Nevertheless, the sequences of both *Laurencia* and *Chondrophycus* clustered next to each other, and these species could belong to the same genus. Similarly, although the two crustose algae that we identified as *Lithothamnion* spp. clustered next to each other, they did not cluster with their reference genus but instead with the genus *Lithophyllum*; these *Lithothamnion* spp. could have been *Lithophyllum*. These two genera of crustose algae both belong to the order Corallinales, which formed a distinct clade from the other Eurhodophytina (subdivision of Rhodophyta). The red macroalgae *Palmaria palmata*, *Dilsea carnosa* (sample number 2), *Turnerella pennyi*, and *Laurencia* sp. grouped in the wrong lineages (Figure 2).

The Euka02 amplicon resolved Phaeophyceae as a monophyletic lineage where most of the sequences grouped either next to their reference or in the neighboring clade. The exceptions were *Dictyosiphon foeniculaceus*, *Chatopteris plumosa*, and Dictyotaceae *Padina* sp. and *Canistrocarpus crispata* (sample number 2). The other Dictyotaceae samples clustered within the lineage and next to the reference sequences. Brown algae orders clustered separately, with all

380 Ectocarpales and all Laminariales forming a single subclade each. However, Fucales and
381 Dictyotales formed two subclades each (Figure 2). Although a genus reference was provided for
382 only 13 brown algae samples, there were seven additional samples that clustered next to the other
383 samples from the same genera. This was the case of all *Saccharina* and all *Laminaria* samples:
384 although there were no references, these samples clustered together at species levels, indicating
385 high barcode discrimination.

386 In contrast, the lineage of green macroalgae (Chlorophyta) displayed taxonomic ambiguity. All
387 reference sequences formed a monophyletic group separate from our sequenced green algal
388 species, despite belonging to the same class (Ulvophyceae, Figure 2). Nevertheless, there was a
389 certain level of identification since the sequenced samples formed a monophyletic group and the
390 order Bryopsidales and most Ulvales separated into their own clades. Only *Valonia ventricosa*
391 clustered within the wrong lineage (among Embryophyta).

392 In some cases, the Euka02 amplicon failed to align sequence variants with the correct
393 taxonomical lineage. When this happened for taxa supported by duplicated samples (two
394 individuals from the same species), and one of the samples was correctly assigned while the other
395 was not, we assumed the primer to be still appropriate for that taxon. This was the case for
396 Phaeophyceae *Padina* sp. and *Canistrocarpus crispata*, Rhodophyta *Dilsea carnosa* and
397 *Laurencia* sp., and the seagrass *Thalassia hemprichii*. In contrast, for unique samples of
398 *Dictyosiphon foeniculaceus* and *Chaetopteris plumosa* (Phaeophyceae), and *Palmaria palmata*
399 and *Turnerella pennyi* (Rhodophyta), we had no reference sequence available for comparison nor
400 duplicated sample. Hence, in these cases, we cannot conclude whether the sequence was wrongly
401 assigned or the primer was not appropriate for these species (Figure 2). Despite those cases,
402 barcoding using the Euka02 primer from the universal 18S gene clearly distinguished each

macroalga lineage and provided the highest taxonomic resolution from a broader set of species.
No other non 18S gene barcode tested here matched these results.

eDNA metabarcoding

The Euka02 primer from the universal 18S gene fingerprinted several lineages of eukaryotic organisms in environmental samples from coastal sediments. Marine macrophytes represented 13.3% of the total reads (362,161 out of 2,727,626); intracellular eDNA contributed the most (72.4% of the 362,161 marine macrophyte reads; Supplementary Table 5). The other taxa included land embryophytes (0.8%), fungi (3.6%), microalgae and other unicellular organisms (14.5% and 0.8%, respectively), and metazoans (arthropods 7.3%, non-arthropods marine invertebrate 36.3%, marine vertebrates 5.2%, vertebrates 0.2%). Unassigned sequences represented 18% of the total reads (Supplementary Table 5).

The resolution of marine macrophyte eDNA taxonomic identification beyond family level increased from 36% to 76% after inclusion of the created DNA reference library from macrophyte tissue. The Euka02 barcode identified 21 marine macrophytes, 14 at genus or species level and the rest at lower taxonomic resolution (Table 5). Five of the seven species present in the sampling sites were identified in the sediments: the seagrasses *Halophila stipulacea*, *H. ovalis*, *Thalassodendron ciliatum* and *Halodule uninervis*, and the mangrove *Avicennia marina*. Nevertheless, the other two species were not available in the created DNA reference library: tissue of *Halophila decipiens* did not amplify with the Euka02 primer, and the sequence of *Thalassia hemprichii* was aligned to the wrong phylum (see Figure 2, within Rhodophyta). *Avicennia marina* was the only mangrove species reported *in situ* and contributed 94.8% of the reads in mangrove sediments; 4.2% of the reads were macroalgae and 0.95% were seagrasses (Table 5; Figure 3). In sediments from seagrass

meadows, mangrove contributed 66.5%, seagrasses 30.4% and macroalgae 3% of the eDNA reads (Table 5; Figure 3). The remaining 16 marine macrophytes from both sediment types were macroalgae: three Chlorophyta, six Phaeophyceae and seven Rhodophyta; of those, eight species are not included in the DNA reference library and were identified at family or order level (Table 5).

Discussion

18S barcoding distinguishes major lineages of marine macrophytes

Out of 18 candidates, the 18S gene primers Euka02 and 18S2 differentiate among macroalgae lineages and marine embryophytes (seagrasses and mangroves), becoming good candidates for barcoding of marine macrophytes. Additionally, the short length of the amplicons makes these primers suitable for metabarcoding of environmental samples, as demonstrated here by fingerprinting marine macrophyte eDNA from coastal sediments.

Depending on the application, barcoding is a trade-off between assigning either few taxa to species level, or many taxa to higher taxonomical levels such as order or even phylum. While other DNA regions may offer higher species-discrimination power, none of the primer pairs tested here amplified a broader set of non-related photosynthetic eukaryotes than those from the 18S region.

A major drawback of the 18S barcoding was the taxonomical ambiguity for green algae. Although the samples did not cluster in the same clade as the reference sequences, 94% of Chlorophyta clustered in a single clade and as a distinct lineage apart from the other macroalgae and embryophytes; most taxa from the same order clustered together and the genera clustered next to each other. This result indicates an acceptable level of identification, yet barcoding of green algae is a challenge and needs improvement. Earlier studies have noted poor performance of both plasmid and mitochondrial markers in some orders of the class Ulvophyceae (to which all our

samples belong) and within the genus *Caulerpa* (same genus for five species tested here) (Famà, Wysor, Kooistra, & Zuccarello, 2002; Kazi, Reddy, & Jha, 2013; Saunders & McDevit, 2012; Zuccarello, Price, Verbruggen, & Leliaert, 2009). The possible high incidence of introns within green algae may complicate the barcoding of Chlorophyta (Saunders & Kucera, 2010; Saunders & McDevit, 2012).

Limitations of plant barcodes for marine macrophyte barcoding

We tested whether the available barcodes for Embryophyta could distinguish among marine macrophytes. The ITS2 locus did not amplify any of the marine macrophytes tested here, in contrast to previous results for seagrass species (Chen et al., 2010; Uchimura, Jean Faye, Shimada, Inoue, & Nakamura, 2008). The trnL primers inconsistently amplified our angiosperms and were unsuccessful in our targeted marine macrophytes. Several studies have shown the usefulness of barcoding regions matK and rbcL on angiosperms, including mangrove and seagrass species (CBOL, 2009; Lucas, Thangaradjou, & Papenbrock, 2012; Newmaster, Fazekas, & Ragupathy, 2006; Yu et al., 2011). However, the performance of these regions on macroalgae is rather limited and ambiguous, as confirmed here and reported earlier (Reef et al., 2017; Saunders & Kucera, 2010; Saunders & McDevit, 2012). Thus, primers used here from the recommended Embryophyta barcodes ITS2, trnL, matK and rbcL are likely not suitable for studies on marine macrophytes.

Saunders (2005) pioneered efforts for DNA barcoding of macroalgae, and developed a set of primers that amplified the COI region in Rhodophyta and Phaeophyceae, while also drawing attention to the low performance of this region for Chlorophyta (McDevit & Saunders, 2009; Saunders & McDevit, 2012). However, we obtained a very low performance using those primers:

only 5 out of 45 macroalgal species amplified. Furthermore, these primers are probably too long (>700 bp) for metabarcoding of fragmented eDNA from sediments.

A suitable primer for metabarcoding marine macrophytes

eDNA metabarcoding of marine macrophyte is a valuable technique for marine environmental studies. For instance, past and present biodiversity assessments of plants and other eukaryotes (Epp et al., 2012; Willerslev et al., 2003) could be extended to marine macrophytes. We identified 21 species of marine macrophytes in Red Sea sediments, of which 14 species recovered were not reported *in situ* during the sampling events. These results show the potential of eDNA analysis to fingerprint the most important macrophytes in marine ecosystems. eDNA fingerprinting of marine macrophytes can also help tracing their contribution to marine food webs, which mainly rely on manual identification of gut content and on stable isotope analyses (de la Morinière et al., 2003; Figgenger, Bernardo, & Plotkin, 2019; Tol, Coles, & Congdon, 2016). eDNA fingerprinting of marine macrophytes is also an outstanding tool for estimating the contribution of marine macrophytes to organic carbon in the open ocean (Ortega et al., 2019) and in coastal habitats (Ortega, 2019). Our metabarcoding not only discriminated between seagrass, mangrove and macroalgae lineages in organic carbon sediments, but provided a first approximation to their relative contributions (Figure 3).

Challenges for barcoding of marine macrophytes

Our initial aim was to identify a unique primer pair for marine angiosperms and a unique primer pair for each of the three macroalga lineages, so that we could use methods such as qPCR to estimate their relative abundance in environmental samples. However, creation of those primers

was challenging and we could not find a specific primer from the tested candidates. When designing *de novo* macroalgal primers, we found that the scarcity of reference libraries restricted development of unique primers, as the investigated taxa are largely missing from the databases. This scarcity was confirmed with the *in silico* PCR amplification, where only 40 of our macroalgae amplified while thousands of microalgae –which are also Chlorophyta, Rhodophyta and Phaeophyceae– were recovered in the analysis (Supplementary Table 4).

Taxonomical resolution below family level is still limited for most marine macrophytes with both 18S candidates (Euka02 and 18S2). This limitation will likely be overcome by increasing sequences in the reference databases. Out of 86 macrophyte species sequenced here, only 41 could be compared against a reference sequence at genus level and 13 at species level, while the remaining macrophyte species were absent in the databases and compared only at family or order level. Hence, there is an urgent need to increase studies and efforts to generate DNA resources, including genotyping of a broad set of species from the three lineages of macroalgae. Currently, universal barcodes like the 18S gene seems to be the most suitable option for recovering less-studied groups such as marine macrophytes.

An inconvenience for sequencing of marine macrophytes, particularly seagrasses and macroalgae, is that these organisms live underwater and are colonized by epiphytes growing on their surface (sponges, cnidarians, gastropods, etc.), or by endophytes that can penetrate their cells (bacteria, fungi, etc.). The use of a universal primer that amplifies both macrophytes and their colonizer increases the challenges for marine macrophyte barcoding, as the non-targeted organisms can add significant noise compared to when the amplification is restricted to targeted taxa. Our Sanger sequences reflected this issue, as several sequences matched marine colonizers even after surface sterilization. For instance, the forward sequence of a *Laurencia* sample matched

to *Janczewskia morimotoi* –its closest parasitic algae (Nonomura & West, 1980), while the reverse sequence matched the real sample (Supplementary Table 1). Although Sanger sequencing can be a faster and a cheaper approach for barcoding of a single organism, it requires high-quality amplicons from the targeted organism (Shokralla et al., 2014). We demonstrate that Sanger may not be appropriate for marine macrophytes where the presence of other DNA sources (such as DNA from colonizer organisms) hampered the sequencing of the targeted taxa; this led to poor quality sequences unacceptable for DNA barcoding. Until specific primers are proven to be successful for targeted DNA amplification of these taxa, we recommend high-throughput sequencing for barcoding of marine macrophytes. High-throughput sequencing has been shown to be an effective method for DNA barcoding of targeted taxa (Shokralla et al., 2014) and even at a low cost (Meier, Wong, Srivathsan, & Foo, 2016); it is also an useful tool for genome skimming and generation of new barcode markers (Bista et al., 2018; Liu, Yang, Zhou, & Zhou, 2017; Straub et al., 2012).

Conclusions

Considering the difficulties in designing macroalgal primers, the limited availability of reference databases, and the need to identify widely-separated lineages such as macroalgae from marine angiosperms, we recommend the 18S region as a universal barcode for macrophytes, particularly for eDNA applications. As shown here, this barcode clearly resolves major phylogenetic patterns within macrophytes at order and family level, and can identify those taxa in marine sediment samples. Furthermore, this region has brought high taxonomical resolution to the animal branch of the tree of life, and we also identified metazoans in our eDNA metabarcoding (Supplementary Table 5). A universal barcode such as the 18S provides a parsimonious approach

540 to resolving not just photosynthetic organisms but most marine eukaryotic biodiversity in a single
541 eDNA sediment sample.

Acknowledgements

This research was supported by King Abdullah University of Science and Technology through baseline support to CMD. DKJ and SBØ received support from the Independent Research Fund Denmark (8021-00222 B, ‘CARMA’). We thank Pierre Taberlet for suggesting the use of primer Euka02 and MiSeq sequencing, and to Craig T. Michell and Chakkiath Paul Antony for their bioinformatic support. We thank Susse Wegeberg, Ole Geertz-Hansen, Karsten Dahl, Peter Stæhr and Michael Bo Rasmussen for contributing with North Atlantic/Arctic species of macroalgae to the barcode library; and Mark A. Tester for identifying terrestrial halophytes from the Red Sea. We thank Wajitha J. Raja Mohamed Sait and Nadia Haj Salah for their help with DNA extractions and sterilization. The authors kindly thank Simor Creer for editing the manuscript and for his willingness and patience to find reviewers; we also thank the three anonymous reviewers for their constructive feedback.

Data accessibility

Reference sequences used in the clusters are available for download from <https://github.com/ngeraldi>. The GenBank accession numbers for the Illumina and Sanger sequences reported in this paper are MK400354-MK400420, MK453128-MK453194, and MT265248-MT265258. Illumina raw sequences are available on NCBI's sequence read archive (SRA) at BioProject PRJNA616121 under BioSample number SAMN14482668. The summary of all barcoding sequences used for the cluster analyses can be found in the Supplementary Information Dataset 1.

Author contributions

AO, NG and CMD conceived the research. AO, RDR, SBØ, and MW conducted the experiments. MW collected and identified Cyprus seagrasses. DKJ provided macroalgae from North Atlantic/Arctic. AO collected and identified macrophytes from the Red Sea, Cyprus, and Finland. AO collected Red Sea sediments with contributions from NG, CMD and DKJ. AO and NG analyzed the data and all coauthors discussed the results. AO wrote the manuscript with input from all authors.

References

- Aires, T., Marbà, N., Serrao, E. A., Duarte, C. M., & Arnaud-Haond, S. (2012). Selective elimination of chloroplastidial DNA for metagenomics of bacteria associated with the green alga *Caulerpa taxifolia* (Bryopsidophyceae) 1. *Journal of Phycology*, 48(2), 483-490.
- Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., & Huse, S. M. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PloS One*, 4(7), e6372.
- Aylagas, E., Borja, Á., Muxika, I., & Rodríguez-Ezpeleta, N. (2018). Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. *Ecological Indicators*, 95, 194-202.
- Baker, P., Minzlaff, U., Schoenle, A., Schwabe, E., Hohlfield, M., Jeuck, A., . . . Brix, S. (2018). Potential contribution of surface-dwelling *Sargassum* algae to deep-sea ecosystems in the southern North Atlantic. *Deep Sea Research (Part II, Topical Studies in Oceanography)*, 148, 21-34.
- Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science & Technology*, 48(3), 1819-1827. doi:10.1021/es404734p
- Bista, I., Carvalho, G. R., Tang, M., Walsh, K., Zhou, X., Hajibabaei, M., . . . Creer, S. (2018). Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. *Molecular Ecology Resources*, 18(5), 1020-1034. doi:10.1111/1755-0998.12888
- Burki, F., Roger, A. J., Brown, M. W., & Simpson, A. G. B. (2019). The New Tree of Eukaryotes. *Trends in Ecology & Evolution*. doi:https://doi.org/10.1016/j.tree.2019.08.008
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581.
- CBOL, P. W. G. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31), 12794-12797.
- Chase, M. W., Cowan, R. S., Hollingsworth, P. M., van den Berg, C., Madriñán, S., Petersen, G., . . . Carine, M. (2007). A proposal for a standardised protocol to barcode all land plants. *Taxon*, 56(2), 295-299.
- Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., . . . Leon, C. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PloS One*, 5(1), e8613. doi:10.1371/journal.pone.0008613
- Coombs, J. T., & Franco, C. M. M. (2003). Isolation and identification of actinobacteria from surface-sterilized wheat roots. *Applied and Environmental Microbiology*, 69(9), 5603-5608.
- De Groot, G. A., During, H. J., Maas, J. W., Schneider, H., Vogel, J. C., & Erkens, R. H. J. (2011). Use of rbcL and trnL-F as a two-locus DNA barcode for identification of NW-European ferns: an ecological perspective. *PloS One*, 6(1), e16371.
- de la Morinière, E. C., Pollux, B. J. A., Nagelkerken, I., Hemminga, M. A., Huiskes, A. H. L., & van der Velde, G. (2003). Ontogenetic dietary changes of coral reef fishes in the mangrove-seagrass-reef continuum: stable isotopes and gut-content analysis. *Marine Ecology Progress Series*, 246, 279-289.

609 Du, G., Wu, F., Mao, Y., Guo, S., Xue, H., & Bi, G. (2014). DNA barcoding assessment of green macroalgae in
610 coastal zone around Qingdao, China. *Journal of Ocean University of China*, 13(1), 97-103.

611 Duarte, C. M. (2017). Reviews and syntheses: Hidden forests, the role of vegetated coastal habitats in the ocean
612 carbon budget. *Biogeosciences*, 14, 301-310. doi:https://doi.org/10.5194/bg-14-301-2017

613 Duarte, C. M., & Cebrián, J. (1996). The fate of marine autotrophic production. *Limnology and Oceanography*,
614 41(8), 1758-1766.

615 Duarte, C. M., & Krause-Jensen, D. (2017). Export from seagrass meadows contributes to marine carbon
616 sequestration. *Frontiers in Marine Science*, 4, 13. doi:10.3389/fmars.2017.00013

617 Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., . . . Johnsen, A. (2012). New
618 environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems.
619 *Molecular ecology*, 21(8), 1821-1833.

620 Famà, P., Wysor, B., Kooistra, W. H. C. F., & Zuccarello, G. C. (2002). Molecular phylogeny of the genus *Caulerpa*
621 (*Caulerpales*, *Chlorophyta*) inferred from chloroplast *tufA* gene1. *Journal of Phycology*, 38(5), 1040-1050.

622 Ficetola, G. F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., . . . Pompanon, F. (2010). An in silico
623 approach for the evaluation of DNA barcodes. *BMC Genomics*, 11(1), 434.

624 Figgenger, C., Bernardo, J., & Plotkin, P. T. (2019). Beyond trophic morphology: stable isotopes reveal ubiquitous
625 versatility in marine turtle trophic ecology. *Biological Reviews*, 94(6), 1947-1973. doi:10.1111/brv.12543

626 Geraldi, N. R., Ortega, A., Serrano, O., Macreadie, P. I., Lovelock, C., Krause-Jensen, D., . . . Kaal, J. (2019).
627 Fingerprinting blue carbon: Rationale and tools to determine the source of organic carbon in marine
628 depositional environments. *Frontiers in Marine Science*, 6, 263.

629 Guardiola, M., Uriz, M. J., Taberlet, P., Coissac, E., Wangenstein, O. S., & Turon, X. (2015). Deep-Sea, Deep-
630 Sequencing: metabarcoding extracellular DNA from sediments of marine canyons. *PloS One*, 10(10),
631 e0139633. doi:10.1371/journal.pone.0139633

632 Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and
633 methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0.
634 *Systematic Biology*, 59(3), 307-321.

635 Guiry, M. D. (2013). AlgaeBase. World-wide electronic publication. Retrieved from <http://www.algaebase.org>

636 Hajibabaei, M., Smith, M., Janzen, D. H., Rodriguez, J. J., Whitfield, J. B., & Hebert, P. D. N. (2006). A minimalist
637 barcode can identify a specimen whose DNA is degraded. *Molecular Ecology Resources*, 6(4), 959-964.

638 Hebert, P. D. N., Gregory, T. R., & Savolainen, V. (2005). The promise of DNA barcoding for taxonomy.
639 *Systematic Biology*, 54(5), 852-859. doi:10.1080/10635150500354886

640 Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit
641 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological*
642 *Sciences*, 270(Suppl 1), S96-S99.

643 Heinecke, L., Epp, L. S., Reschke, M., Stoof-Leichsenring, K. R., Mischke, S., Plessen, B., & Herzsuh, U. (2017).
644 Aquatic macrophyte dynamics in Lake Karakul (Eastern Pamir) over the last 29 cal ka revealed by
645 sedimentary ancient DNA and geochemical analyses of macrofossil remains. *Journal of Paleolimnology*,
646 58(3), 403-417. doi:10.1007/s10933-017-9986-7

- 647 Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). Choosing and using a plant DNA barcode. *PLoS One*,
648 6(5), e19254.
- 649 Kazi, M. A., Reddy, C. R. K., & Jha, B. (2013). Molecular phylogeny and barcoding of *Caulerpa* (Bryopsidales)
650 based on the tufA, rbcL, 18S rDNA and ITS rDNA genes. *PloS One*, 8(12), e82438.
- 651 Krause-Jensen, D., & Duarte, C. M. (2016). Substantial role of macroalgae in marine carbon sequestration. *Nature*
652 *Geoscience*, 9(10), 737-742. doi:10.1038/ngeo2790
- 653 Kuo, L.-Y., Li, F.-W., Chiou, W.-L., & Wang, C.-N. (2011). First insights into fern matK phylogeny. *Molecular*
654 *Phylogenetics and Evolution*, 59(3), 556-566.
- 655 Lahaye, R., Van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., . . . Savolainen, V. (2008). DNA
656 barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences of the*
657 *United States of America*, 105(8), 2923-2928.
- 658 Lever, M. A., Torti, A., Eickenbusch, P., Michaud, A. B., Santl-Temkiv, T., & Jorgensen, B. B. (2015). A modular
659 method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental
660 sample types. *Frontiers in Microbiology*, 6, 476. doi:10.3389/fmicb.2015.00476
- 661 Lieske, E., Fiedler, K. E., & Myers, R. F. (2004). *Coral Reef Guide: Red Sea to Gulf of Aden, South Oman;[the*
662 *Definitive Guide to Over 1200 Species of Underwater Life]*: Collins.
- 663 Little, D. P. (2014). A DNA mini-barcode for land plants. *Molecular Ecology Resources*, 14(3), 437-446.
- 664 Liu, S., Yang, C., Zhou, C., & Zhou, X. (2017). Filling reference gaps via assembling DNA barcodes using high-
665 throughput sequencing—moving toward barcoding the world. *GigaScience*, 6(12).
666 doi:10.1093/gigascience/gix104
- 667 Loneragan, N. R., Bunn, S. E., & Kellaway, D. M. (1997). Are mangroves and seagrasses sources of organic carbon
668 for penaeid prawns in a tropical Australian estuary? A multiple stable-isotope study. *Marine Biology*,
669 130(2), 289-300.
- 670 Lucas, C., Thangaradjou, T., & Papenbrock, J. (2012). Development of a DNA barcoding system for seagrasses:
671 successful but not simple. *PloS One*, 7(1), e29987.
- 672 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*,
673 17(1), pp-10.
- 674 Martin-Lescanne, J., Rousseau, F., De Reviers, B., Payri, C., Couloux, A., Cruaud, C., & Le Gall, L. (2010).
675 Phylogenetic analyses of the *Laurencia* complex (Rhodomelaceae, Ceramiales) support recognition of five
676 genera: *Chondrophycus*, *Laurencia*, *Osmundea*, *Palisada* and *Yuzurua* stat. nov. *European Journal of*
677 *Phycology*, 45(1), 51-61. doi:10.1080/09670260903314292
- 678 McDevit, D. C., & Saunders, G. W. (2009). On the utility of DNA barcoding for species differentiation among
679 brown macroalgae (Phaeophyceae) including a novel extraction protocol. *Phycological Research*, 57(2),
680 131-141. doi:10.1111/j.1440-1835.2009.00530.x
- 681 McKee, K., Rogers, K., & Saintilan, N. (2012). Response of salt marsh and mangrove wetlands to changes in
682 atmospheric CO₂, climate, and sea level. In *Global change and the function and distribution of wetlands*
683 (pp. 63-96): Springer.
- 684 Meier, R., Wong, W., Srivathsan, A., & Foo, M. (2016). \$1 DNA barcodes for reconstructing complex phenomes
685 and finding rare species in specimen-rich samples. *Cladistics*, 32(1), 100-110. doi:10.1111/cla.12115

686 Newmaster, S. G., Fazekas, A. J., & Ragupathy, S. (2006). DNA barcoding in land plants: evaluation of rbcL in a
687 multigene tiered approach. *Botany*, 84(3), 335-341.

688 Nonomura, A. M., & West, J. A. (1980). Ultrastructure of the parasite *Janczewskia morimotoi* and its host *Laurencia*
689 *nipponica* (Ceramiales, Rhodophyta). *Journal of Ultrastructure Research*, 73(2), 183-198.
690 doi:[https://doi.org/10.1016/S0022-5320\(80\)90123-9](https://doi.org/10.1016/S0022-5320(80)90123-9)

691 Ortega, A. (2019). *Fingerprinting marine macrophytes in blue carbon habitats*. (Bioscience – Doctoral dissertation).
692 King Abdullah University of Science and Technology, KAUST. Retrieved from
693 <https://repository.kaust.edu.sa/handle/10754/660253>

694 Ortega, A., Geraldi, N. R., Alam, I., Kamau, A. A., Acinas, S. G., Logares, R., . . . Duarte, C. M. (2019). Important
695 contribution of macroalgae to oceanic carbon sequestration. *Nature Geoscience*, 12(9), 748-754.
696 doi:10.1038/s41561-019-0421-8

697 Pedersen, P. M. (2011). *Grønlands havalger*: Epsilon. dk.

698 Pennisi, E. (2007). Wanted: a barcode for plants. *Science*, 318(5848), 190.

699 Queirós, A. M., Stephens, N., Widdicombe, S., Tait, K., McCoy, S. J., Ingels, J., . . . Carnovale, G. (2019).
700 Connected macroalgal-sediment systems: blue carbon and food webs in the deep coastal ocean. *Ecological*
701 *Monographs*, 89(3), e01366.

702 Rambaut, A. (2007). FigTree, a graphical viewer of phylogenetic trees. Retrieved from
703 <http://tree.bio.ed.ac.uk/software/figtree/>

704 Raven, J. A., Johnston, A. M., Kübler, J. E., Korb, R., McInroy, S. G., Handley, L. L., . . . Vanderklift, M. (2002).
705 Mechanistic interpretation of carbon isotope discrimination by marine macroalgae and seagrasses.
706 *Functional Plant Biology*, 29(3), 355-378.

707 Reef, R., Atwood, T. B., Samper-Villarreal, J., Adame, M. F., Sampayo, E. M., & Lovelock, C. E. (2017). Using
708 eDNA to determine the source of organic carbon in seagrass meadows. *Limnology and Oceanography*,
709 62(3), 1254-1265. doi:10.1002/lno.10499

710 Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: inference of new
711 DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, 39(21), e145-e145.

712 Richmond, M. D. (2002). *A field guide to the seashores of Eastern Africa and the Western Indian Ocean Islands*:
713 Stockholm (Sweden) SIDA/UDSM.

714 Rubinoff, D., Cameron, S., & Will, K. (2006). Are plant DNA barcodes a search for the Holy Grail? *Trends in*
715 *Ecology & Evolution*, 21(1), 1-2.

716 Saunders, G. W. (2005). Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for
717 future applications. *Philosophical Transactions of the Royal Society of London. Series B: Biological*
718 *Sciences*, 360(1462), 1879-1888.

719 Saunders, G. W., & Kucera, H. (2010). An evaluation of rbcL, tufA, UPA, LSU and ITS as DNA barcode markers
720 for the marine green macroalgae. *Cryptogamie Algologie*, 31(4), 487.

721 Saunders, G. W., & McDevit, D. C. (2012). Methods for DNA barcoding photosynthetic protists emphasizing the
722 macroalgae and diatoms. In W. J. Kress & D. L. Erickson (Eds.), *DNA Barcodes: Methods and Protocols*
723 (pp. 207-222). Totowa, NJ: Humana Press.

Shimada, T., Aoki, S., Kameda, K., Hazel, J., Reich, K., & Kamezaki, N. (2014). Site fidelity, ontogenetic shift and diet composition of green turtles *Chelonia mydas* in Japan inferred from stable isotope analysis. *Endangered Species Research*, 25(2), 151-164.

Shokralla, S., Gibson, J. F., Nikbakht, H., Janzen, D. H., Hallwachs, W., & Hajibabaei, M. (2014). Next-generation DNA barcoding: using next-generation sequencing to enhance and accelerate DNA barcode capture from single specimens. *Molecular Ecology Resources*, 14(5), 892-901. doi:10.1111/1755-0998.12236

Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., . . . Bunce, M. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, 7(1), 12240.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H. W., & Richards, T. A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology*, 19(s1), 21-31.

Straub, S. C. K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., & Liston, A. (2012). Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American Journal of Botany*, 99(2), 349-364. doi:10.3732/ajb.1100335

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., . . . Willerslev, E. (2007). Power and limitations of the chloroplast trn L (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, 35(3), e14-e14.

Team, R. (2015). RStudio: integrated development for R. Boston, MA: RStudio, Inc. Retrieved from <http://www.rstudio.com>

Thomsen, P. F., Kielgast, J. O. S., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., . . . Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*, 21(11), 2565-2573.

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA—An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, 183, 4-18.

Tol, S. J., Coles, R. G., & Congdon, B. C. (2016). Dugong dugon feeding in tropical Australian seagrass meadows: implications for conservation planning. *PeerJ*, 4, e2194.

Uchimura, M., Jean Faye, E., Shimada, S., Inoue, T., & Nakamura, Y. (2008). A reassessment of *Halophila* species (Hydrocharitaceae) diversity with special reference to Japanese representatives. *Botanica Marina*, 51(4), 258-268.

Wilkinson, S. P., Davy, S. K., Bunce, M., & Stat, M. (2018). Taxonomic identification of environmental DNA with informatic sequence classification trees. *PeerJ PrePrints*.

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., . . . Cooper, A. (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science*, 300(5620), 791-795.

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., . . . Xiao, P. (2010). Use of ITS2 region as the universal DNA barcode for plants and animals. *PloS One*, 5(10), e13102.

Yu, J., Xue, J. H., & Zhou, S. L. (2011). New universal matK primers for DNA barcoding angiosperms. *Journal of Systematics and Evolution*, 49(3), 176-181.

763 Zuccarello, G. C., Price, N., Verbruggen, H., & Leliaert, F. (2009). Analysis of a plastid multigene data set and the
764 phylogenetic position of the marine macroalga *Caulerpa filiformis* (Chlorophyta). *Journal of Phycology*,
765 45(5), 1206-1212.

Tables

Table 1. Primer pairs tested for the identification of marine macrophytes. Primers amplify ITS2, 18S, matK, COI, rbcL, and trnL DNA regions.

Primer pair	Primer sequence (5'-3')	Length (bp)	Tm (°C)	Reference
18S-V4	F - CCAGCASCYGCGGTAATTCC R - ACTTTCGTTCTTGATYRA	400	57-48	Stoeck et al. 2010
18S1 (V4)	F - CCAGCASCYGCGGTAATTCC R - ACTTTCGTTCTTGATYRA	399	58	Stoeck et al. 2010
18S2 (V9)	F - CCAGCASCYGCGGTAATTCC R - CCTTCYGCAGGTTACACCTA	130	58	Stoeck et al. 2010, Amaral-Zettler et al. 2009
18S Euka02 (V7)	F - TTTGTCTGSTTAATTSCG R - CACAGACCTGTTATTGC	133	53	Guardiola et al. 2015
rbcL	F - GCGGGTGTTAAAGAGTACAA R - AGTAGAAGATTCGGCAGCTA	146	59	This paper
minirbcL F52-rbcLB	F - GTTGGATTCAAAGCTGGTGTTA R - AACCYTCTTCAAAAAGGTC	350	53	Little et al. 2014, Reef et al. 2017
sgmatK1	F - TCCCGGATTCCAGATGTTCC R - AGGAACCGAAAGAGTCTTGGA	136	58	This paper
sgmatK2	F - TGTTACCTTGTTTCGTTTTTGGCA R - CTCCGGACTGCCAAAGGATT	142	58	This paper
mgmatK1	F - ACTGCTTGTTGAAACGGTTAA R - ACCCTCCGATATGATTTGA	130	58	This paper
mgmatK2	F - CGCCAGTGGATTCTATGTTT R - TCGCTCTTTTGATTTTCGGAA	128	58	This paper
COI GF1-GR1	F - TCAACAAATCATAAAGATATTGG R - ACTTCTGGATGTCCAAAAAYCA	740	60	Saunders 2005
COI GF1-DR1	F - TCAACAAATCATAAAGATATTGG R - AAAAYCARAATAAATGTTGA	740	50	Saunders 2005
COI GF2-GR1	F - CCAACCAYAAAGATATWGGTAC R - ACTTCTGGATGTCCAAAAAYCA	740	58	Saunders 2005
COI GF2-DR1	F - CCAACCAYAAAGATATWGGTAC R - AAAAYCARAATAAATGTTGA	740	50	Saunders 2005
ITS2	F - ATGCGATACTTGGTGTGAAT R - GACGCTTCTCCAGACTACAAT	162	56	Yao et al. 2010
trnL c-d	F - CGAAATCGGTAGACGCTACG R - GGGGATAGAGGGACTTGAAC	456	55	Taberlet et al. 2006
trnL g-h	F - GGGCAATCCTGAGCCAA R - CCATTGAGTCTCTGCACCTATC	40	55	Taberlet et al. 2006
trnL bryo_P6 -h	F - GATTTCAGGGAACTTAGGTTG R - CCATTGAGTCTCTGCACCTATC	200	55	Epp et al. 2012, Taberlet et al. 2006

771 **Table 2.** Species used in the barcoding test. Questionable identification is denoted as “?” for some species. Abbreviations: NA, North
772 Atlantic; E., Embryophyta.

Lineage	Species	Location	Lineage	Species	Location
Chlorophyta	<i>Caulerpa brownii</i>	Red Sea	Phaeophyceae	<i>Halidrys siliquosa</i>	NA (Arctic)
Chlorophyta	<i>Caulerpa racemosa</i> var. <i>cylindracea</i>	Red Sea	Phaeophyceae	<i>Laminaria digitata</i>	NA (Arctic)
Chlorophyta	<i>Caulerpa racemosa</i> var. <i>lamourouxii</i> ?	Red Sea	Phaeophyceae	<i>Laminaria solidungula</i>	NA (Arctic)
Chlorophyta	<i>Caulerpa serrulata</i>	Red Sea	Phaeophyceae	<i>Padina boergesenii</i> ?	Red Sea
Chlorophyta	<i>Caulerpa serrulata</i> f. <i>spiralis</i>	Red Sea	Phaeophyceae	<i>Padina boryana</i> ?	Red Sea
Chlorophyta	<i>Caulerpa sertularioides</i>	Red Sea	Phaeophyceae	<i>Padina</i> sp.	Red Sea
Chlorophyta	<i>Caulerpa taxifolia</i>	Red Sea	Phaeophyceae	<i>Punctaria glacialis</i>	NA (Arctic)
Chlorophyta	<i>Chaetomorpha linum</i>	NA (Arctic)	Phaeophyceae	<i>Punctaria plantaginea</i>	NA (Arctic)
Chlorophyta	<i>Codium fragile</i>	NA (Arctic)	Phaeophyceae	<i>Pylaiella varia</i>	NA (Arctic)
Chlorophyta	<i>Dictyosphaeria cavernosa</i>	Red Sea	Phaeophyceae	<i>Saccharina latissima</i>	NA (Arctic)
Chlorophyta	<i>Enteromorpha prolifera</i>	NA (Arctic)	Phaeophyceae	<i>Saccharina nigripes</i>	NA (Arctic)
Chlorophyta	<i>Halimeda discoidea</i> ?	Red Sea	Phaeophyceae	<i>Sargassum aspe</i>	Red Sea
Chlorophyta	<i>Halimeda incrassata</i>	Red Sea	Phaeophyceae	<i>Sargassum ilicifolium</i> ?	Red Sea
Chlorophyta	<i>Halimeda macroloba</i> ?	Red Sea	Phaeophyceae	<i>Sargassum mutium</i> ?	Red Sea
Chlorophyta	<i>Halimeda</i> sp.	Red Sea	Phaeophyceae	<i>Sargassum natans</i> ?	Red Sea
Chlorophyta	<i>Ulva lactuca</i>	NA (Arctic)	Phaeophyceae	<i>Scytosiphon complanatus</i>	NA (Arctic)
Chlorophyta	<i>Ulva</i> ? sp.	Red Sea	Phaeophyceae	<i>Stictyosiphon tortilis</i>	NA (Arctic)
Chlorophyta	<i>Valonia aegagropila</i>	Red Sea	Phaeophyceae	<i>Turbinaria ornata</i> var. <i>serrata</i>	Red Sea
Chlorophyta	<i>Valonia ventricosa</i>	Red Sea	Rhodophyta	<i>Chondrophycus</i> ? <i>papillosus</i> ?	Red Sea
Phaeophyceae	<i>Agarum clathratum</i>	NA (Arctic)	Rhodophyta	<i>Chondrus crispus</i>	NA (Arctic)
Phaeophyceae	<i>Alaria esculenta</i>	NA (Arctic)	Rhodophyta	<i>Coccotylus truncatus</i>	NA (Arctic)
Phaeophyceae	<i>Ascophyllum nodosum</i>	NA (Arctic)	Rhodophyta	<i>Delesseria sanguinea</i>	NA (Arctic)
Phaeophyceae	<i>Chaetopteris plumosa</i>	NA (Arctic)	Rhodophyta	<i>Dilsea carnosa</i>	NA (Arctic)
Phaeophyceae	<i>Cystoseira trinodis</i>	Red Sea	Rhodophyta	<i>Furcellaria humbricalis</i>	NA (Arctic)
Phaeophyceae	<i>Desmarestia aculeata</i>	NA (Arctic)	Rhodophyta	<i>Gracilaria</i> sp.	Red Sea
Phaeophyceae	<i>Dictyosiphon foeniculaceus</i>	NA (Arctic)	Rhodophyta	<i>Hypnea cornuta</i>	Red Sea
Phaeophyceae	<i>Canistrocarpus crispata</i> ?	Red Sea	Rhodophyta	<i>Hypnea</i> sp.	Red Sea
Phaeophyceae	<i>Dictyota humifusa</i> ?	Red Sea	Rhodophyta	<i>Jania pumila</i>	Red Sea
Phaeophyceae	<i>Fucus evanescens</i>	NA (Arctic)	Rhodophyta	<i>Laurencia mcdermidiae</i>	Red Sea
Phaeophyceae	<i>Fucus serratus</i>	NA (Arctic)	Rhodophyta	<i>Laurencia mcdermidiae</i> ?	Red Sea
Phaeophyceae	<i>Fucus vesiculosus</i>	NA (Arctic)	Rhodophyta	<i>Laurencia</i> sp1.	Red Sea

Table 2. Continuation. Species used in the barcoding test.

Lineage	Species	Location	Lineage	Species	Location
Rhodophyta	<i>Laurencia?</i> sp2.	Red Sea	E. (Angiosperm)	<i>Aloe vera</i>	Red Sea
Rhodophyta	<i>Laurencia?</i> sp3.	Red Sea	E. (Angiosperm)	<i>Callistemon viminalis</i>	Red Sea
Rhodophyta	<i>Lithothamnion?</i> sp.	Red Sea	E. (Angiosperm)	<i>Caryota mitis</i>	Red Sea
Rhodophyta	<i>Membranoptera alata</i>	NA (Arctic)	E. (Angiosperm)	<i>Chamaerops hormilis</i>	Red Sea
Rhodophyta	<i>Odonthalia dentata</i>	NA (Arctic)	E. (Angiosperm)	<i>Cocos nucifera</i>	Red Sea
Rhodophyta	<i>Palmaria palmata</i>	NA (Arctic)	E. (Angiosperm)	<i>Cordia subcordata</i>	Red Sea
Rhodophyta	<i>Phycodrys rubens</i>	NA (Arctic)	E. (Angiosperm)	<i>Cymbopogon</i> sp.	Red Sea
Rhodophyta	<i>Phyllophora pseudoceranoides</i>	NA (Arctic)	E. (Angiosperm)	<i>Erythrina crista-galli</i>	Red Sea
Rhodophyta	<i>Polysiphonia arctica</i>	NA (Arctic)	E. (Angiosperm)	<i>Erythrina variegata</i>	Red Sea
Rhodophyta	<i>Polysiphonia nigrescens</i>	NA (Arctic)	E. (Angiosperm)	<i>Eucalyptus</i> sp.	Cyprus
Rhodophyta	<i>Rhodomela confervoides</i>	NA (Arctic)	E. (Angiosperm)	<i>Ficus nitida</i>	Red Sea
Rhodophyta	<i>Turnerella pennyi</i>	NA (Arctic)	E. (Angiosperm)	<i>Halocnemum strobilaceum</i>	Red Sea
Rhodophyta	Unknown sp5. (Rhodophyta)	Red Sea	E. (Angiosperm)	<i>Nerium oleander</i>	Red Sea
Rhodophyta	Unknown sp6. (Rhodophyta)	Red Sea	E. (Angiosperm)	<i>Olea europaea</i>	Cyprus
Macroalgae	Unknown sp1.	Red Sea	E. (Angiosperm)	<i>Phoenix canarensis</i>	Red Sea
Macroalgae	Unknown sp2.	Red Sea	E. (Angiosperm)	<i>Pistacia lentiscus</i>	Red Sea
Macroalgae	Unknown sp3.	Red Sea	E. (Angiosperm)	<i>Plumeria rubra</i>	Red Sea
Macroalgae	Unknown sp4.	Red Sea	E. (Angiosperm)	<i>Punica granatum</i>	Cyprus
E. (Mangrove)	<i>Avicennia marina</i>	Red Sea	E. (Angiosperm)	<i>Ravenala madagascariensis</i>	Red Sea
E. (Mangrove)	<i>Rhizophora mucronata</i>	Red Sea	E. (Angiosperm)	<i>Suaeda monoica</i>	Red Sea
E. (Seagrass)	<i>Cymodocea nodosa</i>	Cyprus	E. (Angiosperm)	<i>Cascabela thevetia</i>	Red Sea
E. (Seagrass)	<i>Cymodocea rotundata</i>	Red Sea	E. (Angiosperm)	<i>Vitis</i> sp.	Cyprus
E. (Seagrass)	<i>Cymodocea serrulata</i>	Red Sea	E. (Angiosperm)	<i>Ziziphus jujuba</i>	Red Sea
E. (Seagrass)	<i>Enhalus acoroides</i>	Red Sea	E. (Other)	<i>Araucaria</i> sp.	Cyprus
E. (Seagrass)	<i>Halodule uninervis</i>	Red Sea	E. (Other)	<i>Callitris</i> sp.	Cyprus
E. (Seagrass)	<i>Halophila decipiens</i>	Red Sea	E. (Other)	<i>Cupressus sempervirens</i>	Cyprus
E. (Seagrass)	<i>Halophila ovalis</i>	Red Sea	E. (Other)	<i>Cycas revoluta</i>	Red Sea
E. (Seagrass)	<i>Halophila stipulacea</i>	Red Sea	E. (Other)	<i>Cyperus</i> sp.	Red Sea
E. (Seagrass)	<i>Posidonia oceanica</i>	Cyprus	E. (Other)	<i>Pinus silvestris</i>	Finland
E. (Seagrass)	<i>Thalassia hemprichii</i>	Red Sea	E. (Other)	<i>Pinus</i> sp.	Finland
E. (Seagrass)	<i>Thalassodendrom ciliatum</i>	Red Sea	E. (Other)	<i>Polypodium vulgare</i>	Finland
E. (Angiosperm)	<i>Adenium obesum</i>	Red Sea	E. (Other)	<i>Pteridium aquilinum</i>	Finland

Table 3. Amplification performance of primer pairs for lineage of macrophytes. Number of samples varied among primers and lineages. Percentage represents species (not total of samples) that amplified among the total number of species tested. Not all amplicons were sequenced. ‘Land other’ includes coniferous trees, ferns, and a cycad species.

Primer pair	Total	Macroalgae	Mangrove	Seagrass	Land angiosperms	Land other
18S2	100% (119/119)	100% (81/81)	100% (2/2)	100% (7/7)	100% (24/24)	100% (5/5)
Euka02	94% (81/86)	97% (64/66)	100% (2/2)	100% (11/11)	100% (2/2)	40% (2/5)
18S-V4	67% (10/15)	63% (5/8)	100% (1/1)	100% (1/1)	60% (3/5)	-
18S1	30% (20/67)	12% (5/42)	50% (1/2)	75% (3/4)	58% (11/19)	-
rbcL	60% (9/15)	67% (4/6)	50% (1/2)	25% (1/4)	100% (3/3)	-
minirbcL	47% (14/30)	50% (7/14)	50% (1/2)	71% (5/7)	0% (0/6)	0% (0/1)
sgmatK2	85% (35/41)	57% (8/14)	100% (2/2)	83% (5/6)	100% (19/19)	-
sgmatK1	21% (7/33)	0% (0/7)	50% (1/2)	50% (2/4)	21% (4/19)	100% (1/1)
mgmatK1	29% (4/14)	0% (0/7)	50% (1/2)	0% (0/2)	100% (3/3)	-
mgmatK2	29% (4/14)	0% (0/7)	50% (1/2)	0% (0/2)	100% (3/3)	-
COI GF1-GR1	7% (3/45)	8% (3/39)	0% (0/1)	0% (0/2)	0% (0/3)	-
COI GF1-DR1	11% (5/45)	13% (5/39)	0% (0/1)	0% (0/2)	0% (0/3)	-
COI GF2-GR1	9% (4/45)	10% (4/39)	0% (0/1)	0% (0/2)	0% (0/3)	-
COI GF2-DR1	2% (1/45)	3% (1/39)	0% (0/1)	0% (0/2)	0% (0/3)	-
ITS2	29% (4/14)	0% (0/7)	0% (0/1)	0% (0/1)	80% (4/5)	-
trnL c-d	20% (2/10)	0% (0/3)	0% (0/1)	0% (0/3)	67% (2/3)	-
trnL g-h	10% (1/10)	0% (0/3)	0% (0/1)	0% (0/3)	33% (1/3)	-
trnL bryo_P6 -h	10% (1/10)	0% (0/3)	0% (0/1)	0% (0/3)	33% (1/3)	-

Table 4. Summary of taxonomic assignment by the two 18S barcode primers Euka02 and 18S2. Total indicates the percentage of assignment to the taxonomic rank among the total number of samples (n = 86 for Euka02 and n= 11 for 18S2). Ref indicates the percentage of sample assignments within the subset of samples that had a reference sequences at that taxonomic level (*n* Ref). For example, the primer Euka02 accurately identified in the right genus 71% of the samples with available genus reference (29 out of 41).

Rank	Euka02			18S2		
	Total	Ref	<i>n</i> Ref	Total	Ref	<i>n</i> Ref
Phylum	88%	88%	86	91%	91%	11
Order	76%	76%	59	45%	45%	11
Family	54%	54%	42	45%	45%	11
Genus	37%	71%	41	27%	30%	10
Species	4%	23%	13	0%	0%	5

795 **Table 5.** List of marine macrophytes identified with the 18S gene mini-barcode using the primer
796 Euka02. All seagrass and mangrove species listed below were spotted on site during the sediment
797 sampling.

Lineage	Species	eDNA Reads
Chlorophyta	<i>Caulerpa serrulata</i> f. <i>spiralis</i>	110
Chlorophyta	<i>Halimeda macroloba</i>	10,551
Phaeophyceae	Ectocarpales	59
Phaeophyceae	<i>Fucus vesiculosus</i>	1,100
Phaeophyceae	Laminariales	274
Phaeophyceae	<i>Padina</i> sp.	226
Phaeophyceae	<i>Saccharina japonica</i>	15
Phaeophyceae	<i>Sargassum</i> sp.	1,109
Rhodophyta	Ceramiaceae	76
Rhodophyta	Corallinales	20
Rhodophyta	Florideophyceae	346
Rhodophyta	<i>Laurencia mcdermidiae</i>	981
Rhodophyta	<i>Lithothamnion</i> sp.	4,886
Rhodophyta	<i>Neosiphonia</i> sp2.	6
Rhodophyta	Rhodomelaceae	1,051
Rhodophyta	<i>Rhodophysema elegans</i>	58
Seagrass	<i>Halodule uninervis</i>	17,948
Seagrass	<i>Halophila ovalis</i>	26,890
Seagrass	<i>Halophila stipulacea</i>	16,822
Seagrass	<i>Thalassodendron ciliatum</i>	10,675
Mangrove	<i>Avicennia marina</i>	467,728
Total		560,931
<i>In situ</i> marine macrophytes		385,629
% <i>in situ</i> marine macrophytes		68.75

Figure Legends

Figure 1. Schematic representation of the steps followed to choose a suitable DNA barcoding to distinguish marine macrophyte lineages.

Figure 2. Taxonomical assignment based on sequence cluster of the macrophytes amplified with the Euka02 primer. The cluster clearly differentiates macrophyte lineages, as most species were accurately assigned to their lineage: 95% of Embryophyta (blue), 94% of Chlorophyta (green), 86% of Phaeophyceae (brown) and 81% of Rhodophyta samples. The reference sequences are labelled with “Ref_” before the species name. Samples from species *Caulerpa serrulata* and *Fucus vesiculosus* have the letter “A” or “B”, indicating that two sequence variants from that sample were chosen for the cluster.

Figure 3. Location of the mangrove forest (tree icon) and seagrass meadows (grass icon) where sediments were sampled in the Red Sea. Pies indicate the proportion of marine macrophytes to the eDNA sediment pool by sampling location. Bars indicate the mean proportion of eDNA from seagrasses, mangroves and macroalgae by coastal habitat.

817 **Supplementary Information**

818 **Supplementary Table Legends (Tables are in the SI_Supplementary_Tables.xlsx file)**

819 **Supplementary Table 1.** Summary of Sanger sequencing attempts. BLAST match indicates the
820 taxon to which the sequence was assigned.

821
822 **Supplementary Table 2.** Amplification performance of the six barcodes and the 18 primers
823 tested. X indicates failed amplification, Y indicates successful amplification, - indicates that the
824 primer was not tested on that species.

825
826 **Supplementary Table 3.** Results of *in silico* PCR for seven primer pairs aiming marine
827 macrophyte sequences from the EMBL database; eight mismatches were allowed. Data indicate
828 the taxonomical level outputs (ecotaxspecificity). “Taxon_total” indicates the number of taxa to
829 distinguish, “Taxon_ok” indicates the taxa that can be distinguished based on the database.
830 “Percentage” indicates the identification percentage of each taxa present in the EMBL
831 sequences. “Total algal species” indicates all Chlorophyta, Rhodophyta and Phaeophyceae that
832 amplified *in silico*; those species also include unicellular microalgae.

833
834 **Supplementary Table 4.** *In silico* PCR amplification of macroalgae included in this study, using
835 seven primer pairs. Yes or No indicates whether the species amplified *in silico*. + or - indicates
836 whether the species amplified or not during the lab experiments.

837
838 **Supplementary Table 5.** Eukaryotic diversity based on intracellular and extracellular eDNA
839 from coastal sediment samples, and using the primer Euka02 from the 18S gene. Values
840 represent eDNA reads; *Marine macrophytes* include macroalgae (Chlorophyta, Rhodophyta and
841 Phaeophyceae), mangrove and seagrass.

Supplementary Figure Legend

Supplementary Figure 1. Taxonomical assignment based on sequence cluster of the macrophytes amplified with the 18S2 primer. The cluster clearly differentiates macrophyte lineages. The reference sequences are labelled with “Ref_” before the species name. Some species names have the letter “A”, “B” or “C”, indicating that two or three sequence variants from that sample were chosen for the cluster. Contrary to primer Euka02, 18S2 primer clustered *Valonia ventricosa* next to its reference, and two *Halimeda* species next to the reference sequences of the class.

Supplementary Figure 2. Evidence of the amplification of eDNA from marine sediments using the 18S2 primer pairs. Upper gel: amplification of eDNA from Red Sea sediments. Lower gel: amplification of eDNA from Arctic marine sediments.

Step-by-step protocol for intracellular and extracellular environmental DNA extraction based on Lever et al. 2015

Instruments

- HulaMixer (Invitrogen, Merelbeke, Belgium)
- Thermomixer (Eppendorf, Hamburg, Germany)
- Vortex
- Centrifuges for 15 ml and 2 ml tubes (14.000 x g)
- Vacuum centrifuge

Notes before starting

1. Prepare reagents following Lever et al. 2015
2. Make sure all materials for reagents and for extraction are sterilized with best available method (i.e. autoclave, UV light, DNA killer)
3. Regularly clean gloves with DNA killer throughout the extraction
4. Take every precaution to reduce cross-contamination
5. Some tips are listed in the Notes.

Procedure for extraction of intracellular and extracellular eDNA

1. Weigh 0.2 g of sediment in a 15 ml tube.
Note: weight empty tubes before. Thaw and slightly shake sediment samples for easy transfer to the new tubes. Add an extra tube to run the extraction control.
2. Sorption prevention of nucleic acids
 - a) Add 0.5 ml $\text{PO}_4 \text{ g}^{-1}$
Note: pH should be >9 , add NaOH until desired pH is reached. Use 10-100 $\mu\text{mol PO}_4 \text{ g}^{-1}$ for organic-rich and sandy sediments, and 100-1000 $\mu\text{mol PO}_4 \text{ g}^{-1}$ for carbonate-rich sediments.
 - b) Mix gently to coat the sample with PO_4 . **Note:** it works fine for 5 min, 250 rpm.
 - c) Add 9 volumes of 10x TE buffer
 - d) Gently shake in the HulaMixer at room temperature for 2 hours
 - e) Centrifuge for 20 min at 10,000 x g. **Note:** meanwhile, set up syringes and filters.
Optional for carbonate-rich samples:
 - f) Add 1 volume of CDM to the sample
 - g) Gently shake at room temperature for 1 hour
 - h) Add 8 volumes of 10x TE buffer
 - i) Gently shake at room temperature for 2 hours
 - j) Centrifuge for 20 min at 10,000 x g. **Note:** meanwhile, set up syringes and filters.
3. Extracellular and intracellular eDNA separation
 - a) Transfer filtered supernatant (extracellular eDNA) to a 15 ml tube, use 0.2 μm filter attached to a 5 ml syringe, proceed to first precipitation (step 4)
 - b) Keep sediment (intracellular eDNA) frozen while proceeding to cell lysis (step 8)
4. First precipitation
 - a) Add LPA linear polyacrylamide (final concentration 20 μl per ml of sample)
 - b) Add 2 volumes of PEG 6000 (30% w/v with 1.6 M NaCl)
 - c) Homogenize by hand shaking for 30 sec, or vortex for a few seconds

- 902 d) Incubate at room temperature in dark for 2 hours
903 **Note:** safe to stop here. Samples can be incubated overnight at 4°C.
904 e) Centrifuge at 14,000 x g for 60 min. **Note:** meanwhile, label a batch of 2 ml tubes.
905 f) Carefully discard the supernatant and keep pellet (pipette off or use vacuum)
906 **Note:** leave ~50 µl of supernatant to avoid pellet lost.
907 g) Add 1 ml of 70% ethanol to the 15 ml tube, mix by pipetting
908 h) Transfer the sample into a 2 ml tube
909 i) Centrifuge 3 min at 14,000 x g, discard supernatant
910 j) Repeat ethanol addition (1 ml), centrifuge (5 min), discard supernatant
911 k) Dry pellet using vacuum centrifuge or oven (45°C, 15-40 min). Do not over dry
912 l) Dissolve pellet in 100 µl water
913 5. Clean-up of extracellular DNA
914 a) Add 500 µl of lysis solution I
915 b) Gently shake for 1 hour at 50°C (350 rpm)
916 **Note:** meanwhile, label a new batch of 2 ml tubes.
917 c) **For organic-rich sediment** add 500 µl of lysis solution II and incubate a second
918 time for 1 hour at 65°C, otherwise skip this step
919 d) Centrifuge for 10 min at 4°C and 10,000 x g
920 e) Transfer supernatant to a clean tube
921 6. Purification (perform in a fume hood)
922 a) Add 1 volume (600 µl) of Chloroform-isoamyl alcohol (CI)
923 b) Vortex at max speed for 10 sec
924 c) Centrifuge for 10 min at 4°C and 10,000 x g.
925 **Note:** meanwhile, label a new batch of 2 ml tubes.
926 d) Transfer supernatant into a clean tube
927 **Note:** take top aqueous layer and leave precipitants in between layers.
928 e) Repeat CI addition, vortex and centrifugation
929 7. Second Precipitation
930 a) Add LPA linear polyacrylamide (final concentration 20 µl per ml of sample)
931 b) Add 2 volumes of PEG 6000 (30% w/v with 1.6 M NaCl)
932 c) Homogenize by hand shaking for 30 sec, or vortex for a few seconds
933 d) Incubate at room temperature in dark for 2 hours
934 **Note:** safe to stop here. Samples can be incubated overnight at 4°C.
935 e) Centrifuge at 14,000 x g for 10 min. **Note:** label a new batch of 2 ml tubes.
936 f) Carefully discard the supernatant and keep pellet (pipette off or use vacuum)
937 **Note:** leave ~50 µl of supernatant to avoid pellet lost.
938 g) Add 650 µl of 70% ethanol, mix by pipetting
939 h) Transfer the sample into a 2 ml tube
940 i) Centrifuge 2 min at 14,000 x g, discard supernatant
941 j) Repeat ethanol addition (1 ml), centrifuge (3 min), discard supernatant
942 k) Dry pellet using vacuum centrifuge or oven (45°C, 15-20 min). Do not over dry
943 l) Dissolve pellet in 100 µl PCR water or DNA elution buffer
944 Extracellular eDNA is ready for PCR
945

946 Continuation of intracellular eDNA extraction (from 3b)

- 947
948 8. Cell lysis, starting on frozen sediment in 15 ml tube
949 a) Add 2.5 volumes of lysis solution I (500 µl lysis solution I per 0.2 g sample)
950 b) Gently homogenize, then freeze sample at -80°C 5 min
951 **Follow one of the below lysis protocols (LP; perform in fume hood)**
952 I. LP1 (quick, suitable for most high-throughput applications and RNA)
953 - Add 2.5 volumes of cold PCI (500 µl per 0.2 g sample)
954 - Thaw sample
955 - Vortex for 10 sec at high speed
956 - Disrupt sediment using beads at medium-high speed in a TissueLyser (0.5-1
957 min)
958 II. LP2 (most versatile)
959 - Thaw sample
960 - Vortex for 10 sec at high speed
961 - Shake for 1 hour at 50°C
962 - Disrupt sediment using beads at medium-high speed in a TissueLyser (0.5-1
963 min), or vortex for 10 min
964 **Optional:** to increase yield repeat cycle of freezing, thawing and heat-shaking
965 the sample a second time.
966 III. LP3 (best for organic rich samples)
967 - Thaw sample
968 - Vortex for 10 sec at high speed
969 - Shake for 1 hour at 50°C
970 - Repeat cycle of freezing at -80°C for 5 min, thawing and heat-shaking
971 - Add 2.5 volumes (500 µl per 0.2 g sample) of Lysis solution II
972 - Repeat cycle of freezing at -80°C for 5 min, thawing and heat-shaking
973 - Repeat a second time the cycle of freezing at -80°C for 5 min, thawing and
974 heat-shaking (this time 1 hour at 60°C)
975 c) Centrifuge for 15 min at 10,000 x g and 4°C
976 d) Transfer supernatant to a clean tube, discard pellet and flow-through
977
978 9. Follow extracellular protocol starting on Purification (step 6). Intracellular DNA is ready
979 for PCR after Second Precipitation (step 7). Tip: test for best PCR performance with
980 eDNA dilutions at 1:1, 1:10, 1:50 or 1:100; dilutions reduce PCR inhibitors carried out
981 from the sediment.